

SEQUENCE LISTING

5

(1) GENERAL INFORMATION

(i) APPLICANT: MICHALOVICH, DAVID
HAYES, PHILIP DAVID

10

(ii) TITLE OF THE INVENTION: NOVEL COMPOUNDS

(iii) NUMBER OF SEQUENCES: 4

15

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Ratner & Prestia
(B) STREET: P.O. Box 980
(C) CITY: Valley Forge
(D) STATE: PA
(E) COUNTRY: USA
(F) ZIP: 19482

20

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

25

30

T, 260

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: TO BE ASSIGNED
(B) FILING DATE: 29-JAN-1999
(C) CLASSIFICATION: UNKNOWN

35

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: UK APPLICATION NO. TO BE ASSIGNED
(B) FILING DATE: 20-JAN-1999

40

(A) APPLICATION NUMBER: EP APPLICATION NO. 98300694.1
(B) FILING DATE: 30-JAN-1998

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Prestia, Paul F
(B) REGISTRATION NUMBER: 23,031
(C) REFERENCE/DOCKET NUMBER: GP-30039

45

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 610-407-0700
(B) TELEFAX: 610-407-0700
(C) TELEX: 846169

50

(2) INFORMATION FOR SEQ ID NO:1:

55

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2010 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

60

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

5 ATGACGCCTC CTCCGCCCGG ACGTGCCGCC CCCAGCGCAC CGCGCGCCCG CGTCCCTGGC 60
 CCGCCGGCTC GGTGCGGGCT TCCGCTGCGG CTGCGGCTGC TGCTGCTGCT CTGGGCGGCC 120
 GCCGCCTCCG CCCAGGGCCA CCTAAGGAGC GGACCCCGCA TCTTCGCCGT CTGGAAAGGC 180
 CATGTAGGGC AGGACCGGGT GGACTTTGGC CAGACTGAGC CGCACACGGT GCTTTTCCAC 240
 GAGCCAGGCA GCTCCTCTGT GTGGGTGGGA GGACGTGGCA AGGTCTACCT CTTTGA CTTC 300
 CCCGAGGGCA AGAACGCATC TGTGCGCACG GTGAATATCG GCTCCACAAA GGGGTCTGT 360
 CTGGATAAGC GGGACTGCGA GAACTACATC ACTCTCCTGG AGAGGCGGAG TGAGGGGCTG 420
 10 CTGGCCTGTG GCACCAACGC CCGGCACCCC AGCTGCTGGA ACCTGGTGAA TGGCACTGTG 480
 GTGCCACTTG GCGAGATGAG AGGCTACGCC CCCTTCAGCC CGGACGAGAA CTCCCTGGTT 540
 CTGTTTGAAG GGGACGAGGT GTATTCCACC ATCCCGAAGC AGGAATACAA TGGGAAGATC 600
 CCTCGGTTCC GCCGCATCCG GGGCGAGAGT GAGCTGTACA CCAGTGATAC TGTCTATGCAG 660
 AACCACAGT TCATCAAAGC CACCATCTGT CACCAAGACC AGGCTTACGA TGACAAGATC 720
 15 TACTACTTCT TCCGAGAGGA CAATCCTGAC AAGAATCTTG AGGCTCCTCT CAATGTGTCC 780
 CGTGTGGCCC AGTTGTGCAG GGGGGACCAG GGTGGGGAAA GTTCACTGTC AGTCTCCAAG 840
 TGGAACACTT TTCTGAAAGC CATGCTGGTA TGCAGTGATG CTGCCACCAA CAAGAACTTC 900
 AACAGGCTGC AAGACGTCTT CCTGCTCCCT GACCCAGCG GCCAGTGGAG GGACACCAGG 960
 GTCTATGGTG TTTTCTCCAA CCCCTGGAAC TACTCAGCCG TCTGTGTGTA TTCCCTCGGT 1020
 20 GACATTGACA AGGTCTTCCG TACCTCCTCA CTCGAGGGCT ACCACTCAAG CCTTCCCAAC 1080
 CCGCGGCCTG GCAAGTGCCT CCCAGACCAG CAGCCGATAC CCACAGAGAC CTTCCAGGTG 1140
 GCTGACCGTC ACCCAGAGGT GGCGCAGAGG GTGGAGCCCA TGGGGCCTCT GAAGACGCCA 1200
 TTGTTCCACT CTAAATACCA CTACCAGAAA GTGGCCGTCC ACCGCATGCA AGCCAGCCAC 1260
 GGGGAGACCT TTCATGTGCT TTACCTAAT ACAGACAGGG GCACTATCCA CAAGGTGGTG 1320
 25 GAACCGGGGG AGCAGGAGCA CAGCTTCGCC TTCAACATCA TGGAGATCCA GCCCTTCCGC 1380
 CGCGCGGCTG CCATCCAGAC CATGTCGCTG GATGCTGAGC GGAGGAAGCT GTATGTGAGC 1440
 TCCCAGTGGG AGGTGAGCCA GGTGCCCTG GACCTGTGTG AGGTCTATGG CGGGGGCTGC 1500
 CACGGTTGCC TCATGTCCCG AGACCCCTAC TGCGGCTGGG ACCAAGGCCG CTGCATCTCC 1560
 ATCTACAGT CCGAACGGTC AGTGCTGCAA TCCATTAATC CAGCCGAGCC ACACAAGGAG 1620
 30 TGTCCCAACC CCAAACCAGA CAAGCCCCA CTGCAGAAGG TTTCCCTGGC CCCAACTCT 1680
 CGCTACTACC TGAGCTGCCC CATGGAATCC CGCCACGCCA CCTACTCATG GCGCCACAAG 1740
 GAGAACGTGG AGCAGAGCTG CGAACCTGGT CACCAGAGCC CCAACTGCAT CCTGTTTCATC 1800
 GAGAACCTCA CGGCGCAGCA GTACGGCCAC TACTTCTGCG AGGCCAGGA GGGCTCCTAC 1860
 35 TTCCGCGAGG CTCAGCACTG GCAGCTGCTG CCCGAGGACG GCATCATGGC CGAGCACCTG 1920
 CTGGGTCATG CCTGTGCCCT GGCCGCCTCC CTCTGGCTGG GGGTGCTGCC CACACTCACT 1980
 CTTGGCTTGC TGGTCCACTA GGGCCTCCCG 2010

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 666 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

50 Met Thr Pro Pro Pro 5 Pro Gly Arg Ala Ala Pro Ser Ala Pro Arg Ala
 1 5 10 15
 Arg Val Pro Gly Pro Pro Ala Arg Leu Gly Leu Pro Leu Arg Leu Arg
 20 25 30
 55 Leu Leu Leu Leu Leu Trp Ala Ala Ala Ala Ser Ala Gln Gly His Leu
 35 40 45
 Arg Ser Gly Pro Arg Ile Phe Ala Val Trp Lys Gly His Val Gly Gln
 50 55 60
 60 Asp Arg Val Asp Phe Gly Gln Thr Glu Pro His Thr Val Leu Phe His
 65 70 75 80
 Glu Pro Gly Ser Ser Ser Val Trp Val Gly Gly Arg Gly Lys Val Tyr
 85 90 95
 Leu Phe Asp Phe Pro Glu Gly Lys Asn Ala Ser Val Arg Thr Val Asn
 100 105 110

5 Ile Gly Ser Thr Lys Gly Ser Cys Leu Asp Lys Arg Asp Cys Glu Asn
 Tyr Ile Thr Leu Leu Glu Arg Arg Ser Glu Gly Leu Leu Ala Cys Gly
 10 Asn Ser Leu Val Leu Phe Glu Gly Asp Glu Val Tyr Ser Thr Ile Arg
 Lys Gln Glu Tyr Asn Gly Lys Ile Pro Arg Phe Arg Arg Ile Arg Gly
 15 Glu Ser Glu Leu Tyr Thr Ser Asp Thr Val Met Gln Asn Pro Gln Phe
 Ile Lys Ala Thr Ile Val His Gln Asp Gln Ala Tyr Asp Asp Lys Ile
 Tyr Tyr Phe Phe Arg Glu Asp Asn Pro Asp Lys Asn Pro Glu Ala Pro
 20 Leu Asn Val Ser Arg Val Ala Gln Leu Cys Arg Gly Asp Gln Gly Gly
 Glu Ser Ser Leu Ser Val Ser Lys Trp Asn Thr Phe Leu Lys Ala Met
 25 Leu Val Cys Ser Asp Ala Ala Thr Asn Lys Asn Phe Asn Arg Leu Gln
 Asp Val Phe Leu Leu Pro Asp Pro Ser Gly Gln Trp Arg Asp Thr Arg
 Val Tyr Gly Val Phe Ser Asn Pro Trp Asn Tyr Ser Ala Val Cys Val
 30 Tyr Ser Leu Gly Asp Ile Asp Lys Val Phe Arg Thr Ser Ser Leu Lys
 Gly Tyr His Ser Ser Leu Pro Asn Pro Arg Pro Gly Lys Cys Leu Pro
 35 Asp Gln Gln Pro Ile Pro Thr Glu Thr Phe Gln Val Ala Asp Arg His
 Pro Glu Val Ala Gln Arg Val Glu Pro Met Gly Pro Leu Lys Thr Pro
 Leu Phe His Ser Lys Tyr His Tyr Gln Lys Val Ala Val His Arg Met
 40 Gln Ala Ser His Gly Glu Thr Phe His Val Leu Tyr Leu Thr Thr Asp
 Arg Gly Thr Ile His Lys Val Val Glu Pro Gly Glu Gln Glu His Ser
 Phe Ala Phe Asn Ile Met Glu Ile Gln Pro Phe Arg Arg Ala Ala Ala
 45 Ile Gln Thr Met Ser Leu Asp Ala Glu Arg Arg Lys Leu Tyr Val Ser
 Ser Gln Trp Glu Val Ser Gln Val Pro Leu Asp Leu Cys Glu Val Tyr
 50 Gly Gly Gly Cys His Gly Cys Leu Met Ser Arg Asp Pro Tyr Cys Gly
 Trp Asp Gln Gly Arg Cys Ile Ser Ile Tyr Ser Ser Glu Arg Ser Val
 55 Leu Gln Ser Ile Asn Pro Ala Glu Pro His Lys Glu Cys Pro Asn Pro
 Lys Pro Asp Lys Ala Pro Leu Gln Lys Val Ser Leu Ala Pro Asn Ser
 Arg Tyr Tyr Leu Ser Cys Pro Met Glu Ser Arg His Ala Thr Tyr Ser
 60 Trp Arg His Lys Glu Asn Val Glu Gln Ser Cys Glu Pro Gly His Gln
 Ser Pro Asn Cys Ile Leu Phe Ile Glu Asn Leu Thr Ala Gln Gln Tyr

Gly His Tyr Phe Cys Glu Ala Gln Glu Gly Ser Tyr Phe Arg Glu Ala
 610 615
 5 Gln His Trp Gln Leu Leu Pro Glu Asp Gly Ile Met Ala Glu His Leu
 625 630 635 640
 Leu Gly His Ala Cys Ala Leu Ala Ala Ser Leu Trp Leu Gly Val Leu
 645 650 655
 Pro Thr Leu Thr Leu Gly Leu Leu Val His
 660 665

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 712 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCGCCTGCCG CCCAGGGCCA CCTAAGGAGC GGATNCTANN TCTTCGCCGT CTGGAAAGGC 60
 CATGTAGGGC AGGACCGGGT GGACTTTGGC CAGACTGAGC CGCACACGGT GCTTTTCCAC 120
 25 GAGCCAGGCA GTCCTCTGT GTGGGTGGGA GGACGTGGCA AGGTCTACCT CTTTGA CTTC
 CCCGAGGGCA AGAACGCATC TGTGCGCAGC GTGAATATCG GCTCCACAAA GGGGTCCTGT 240
 CTGGATAAGC GGGACTGCGA GAACTACATC ACTCTCCTGG AGAGGCGGAG TGAGGGGCTG 300
 CTGGCCTGTG GCACCAACGC CCGGCACCCC AGCTGCTGGA ACCTGGTGAA TGCACGTGTG 360
 TGCCACCTTG GCGAGAGTGG AGGCTACGCC CCCTTCAGCC CGGACGAGAA CGTCCCGTGG 420
 30 TTCTGTTTTG AAGGGGACGA AGTGATTCC ACCATCCGGA AAGCAAGGAA TTACAATTGG 480
 GAAGATCCTC GGTTCGCGCC CATCCGGGGC GAGAGTGAGC TGTACACCAG TGATACTGTC 540
 ATGCAGAACC CACAGTTCAT CAAAGCCACC ATCGTGACCC AAGACCAGGC TTACGATGAC 600
 AAGATCTACT ACTTCTTCCG AGAGGACAAT CCTGACAAGA ATCCTGAGGC TCCTCTCAAT 660
 35 GTGTCCCGTG TGGCCAGTGT GTGCAGGGGG GACCAGGGTG GGGAAAGTTC AN 712

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 215 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Gly Gln Asp Arg Val Asp Phe Gly Gln Thr Glu Pro His Thr Val Leu
 1 5 10 15
 50 Phe His Glu Pro Gly Ser Ser Ser Val Trp Val Gly Gly Arg Gly Lys
 20 25 30
 Val Tyr Leu Phe Asp Phe Pro Glu Gly Lys Asn Ala Ser Val Arg Thr
 35 40 45
 55 Val Asn Ile Gly Ser Thr Lys Gly Ser Cys Leu Asp Lys Arg Asp Cys
 50 55 60
 Glu Asn Tyr Ile Thr Leu Leu Glu Arg Arg Ser Glu Gly Leu Leu Ala
 65 70 75 80
 Cys Gly Thr Asn Ala Arg His Pro Ser Cys Trp Asn Leu Val Asn Ala
 85 90 95
 60 Leu Trp Cys His Leu Gly Glu Ser Gly Gly Tyr Ala Pro Phe Ser Pro
 100 105 110
 Asp Glu Asn Val Pro Trp Phe Cys Phe Glu Gly Asp Glu Val Tyr Ser
 115 120 125

15

1. The first part of the report, which is the most important, is the *Summary*. This should be written in a clear, concise, and factual manner. It should state the purpose of the study, the methods used, the results obtained, and the conclusions drawn. It should be written in a way that is easy to read and understand.